Distr.
Limited

E/ECA/STAT/WG/MS/89/9
5 October 1989

Original: English

UNITED NATIONS

ECONOMIC AND SOCIAL COUNCIL

ECONOMIC COMMISSION FOR AFRICA

Working Group Meeting
on Migration Statistics

Addis Ababa, 6-10 November 1989

INTERLINKING MIGRATION DATA FROM DIFFERENT SOURCES

C/NAIROBI.

TABLE OF CONTENTS

		Paragraph	Page
1.	INTRODUCTION	1 - 6	1 - 2
2.	INTERNAL MIGRATION	7 - 33	2
	a) Census and census	8 - 14	2 - 4
	b) Census and civil/vital registration record	15 - 16	4 - 5
	c) Census and survey	17 - 23	5 - 6
	d) Census and population registers	24 - 25	6 - 7
	e) Survey and survey	26 - 33	7 - 9
3.	INTERNATIONAL MIGRATION	34	9 - 11
4.	DUAL RECORD SYSTEM	35 - 37	11 - 11
5.	CONCLUSION	38	12
	ANNEX		

1. INTRODUCTION

1. Actual and potential sources of migration data in Africa are: population censuses, demographic and labour force surveys, population registers and, for international migration only, border control records. Each of these sources by itself may not provide adequate and reliable data for measuring internal as well as international migration. However, by combining two sources of the same kind (e.g. the 1970 and 1980 population census of a country) or by combining data at the macro or micro level from two different sources (e.g. census and survey), it is possible to obtain data which can be used to measure and in certain cases interpret migration levels and trends.

2. Regretably, migration research is one of the neglected areas of demography in Africa. The general view is that the sparsity of migration data is responsible for this state of affairs and that it will need a major social science research project on migration surveys on the lines of the World Fertility Survey or the Demographic and Health Survey to stimulate interest in migration studies.

3. Efforts to conduct subregional migration surveys in Africa have been largely unsuccessful. In 1972, the World Bank, the Organization for Economic Co-operation and Development and a group of African researchers interested in migration research got together and discussed the modalities in organizing such a survey in nine West African States (Benin, Burkina Faso, Ghana, Côte d'Ivoire, Togo, Nigeria, Mali, Niger, Senegal). Later the number was reduced to five by deleting Benin, Togo, Nigeria and Niger. Desk studies on available data were published by K.C. Zechariah, J. Conde et al but the surveys never got off the ground due to operational difficulties. Thus one of the more encouraging attempts at effective collaboration between individual African scholars and multilateral institutions came to nothing.

4. In spite of this, there should have been concerted efforts to follow up on the original studies by Zechariah et al by examining all availabe data not only from population censuses but also from other data sources. This would have helped to enhance the awareness of the policy makers and planners in the West African region to level and trends in migration and their implications.

5. A number of surveys and studies on migration has also been carried out in other subregions, especially in Southern Africa. These studies, however, have not in general attempted to link directly data from various sources to get a clearer picture of migration.

6. In this paper, an attempt will be made to illustrate the various methods of interlinking data from different sources to help improve the stock of data available for measuring and interpreting migration. This will be done separately for internal and international migration.

2. INTERNAL MIGRATION

7. Interlinking internal migration data from different sources involves simple as well as very elaborate manipulation of data sets. In this section, some of the actual methods of interlinking data will be considered. In addition possible new approaches will be outlined.

(a) Census and Census:

8. One of the more traditional methods of interlinking internal migration data from different sources is the use of data from successive population censuses to measure intercensal or period interregional net migration. This involves using data on birth place and place of usual residence. In the absence of place of usual residence data, place of enumeration may be used as proxy. The limitations of this alternative approach are well known and data so obtained should be treated with caution.

9. UN Manual VI (1970) describes three approaches of using data from two successive population censuses to obtain estimates of life-time in-migrants and out-migrants and net intercensal migration. These methods which assume that the population is closed or quasi closed are summarised as follows:

(i) Use of an over-all survival ratio (e.g. ratio of native born persons aged 10 years and over in the 1980 census over native born of all ages enumerated in the 1970 census) to estimate the number of in-and out-migrants from each region or domain of interest, provided the data from the two successive censuses are or can be classified by those regions. The overall life table survival ratio may also be used if the life table for the entire period is available. This approach is used when the age-distribution of the out-born population is not available.

(ii) Compute the survival ratio for persons by sex born in region i and enumerated anywhere in the country (i = 1, 2,.....n), where n is the total number of regions in the country. The survival ratio is obtained by dividing the total number aged 10 years or more born in region i but enumerated anywhere in the country in the second census by the total number of persons born and enumerated in region i in the preceding census. For the successful application of

this method, data in the first census should be by region and sex and for the second, age, sex and region

(iii) Use of the census survival ratio method. This involves calculating for each age cohort and each sex, the survival ratios, namely population aged x + n in the second census as a proportion of population aged x in the first census, separately for each area. As already stated, the computation should be done separately for each sex. The expected number of persons by sex who should have been enumerated in each of the regions can then be obtained by subtracting the expected from the enumerated. The net change due to migration in the period between the two successive censuses e.g. 1970-1980 can be then obtained. For this method, data should be by region and sex and age.

10. It should be noted that the third approach cannot give estimates of the migration of children born during the intercensal period. Also, the use of birth place data in all these methods implies that the usual place of residence of the mother at the time of birth is identical with place of birth. However, in Africa, it is common practice for some women to go to their mother's place of residence to deliver. In such cases, the place of physical birth is not the place of usual residence of the mother. Migration data collected by means of place of birth statistics are thus distorted. Ghana has tried to solve this problem by defining place of birth in its censuses as the place of usual residence of the mother at the time of birth (of the child). However, no evaluation of the effect of this definition on the quality and usefulness of migration data collected has been made.

11. The calculations above are generally valid if the population is closed or approximately closed to international migration. This is a necessary condition for the application of a survival ratio method.

12. Other types of analysis are also possibility based on different types of models. In the United Nations publication, Towards a System of Social and Demographic Statistics an illustration of the type of analysis based on a demographic matrix for several regions is given. A generalised form of this matrix for the purposes of analysing migration data based on two successive censuses could be stated as in the following example:

1970	Other world (births, 1970 - 1980)			Region	Rest of country	Other countries	Totals
1980	Region i	Rest of country	Other countries				
Other world Deaths, (1970 - 1980)							
Region i							
Rest of country							
Other countries							
Total							

13. By following procedures outlined in the document under reference it is possible to estimate the proportions of the different age groups of the population moving from region i to region j during the intercensal period. In the example of Oliver cited in the publication under consideration, the data tended to suggest that "migration may be more effective in reducing disparities of unemployment in bad times than in good and, also, that over the period its influence in this respect tended to decline".

14. The above example has been cited to show that in-depth analysis of migration data is possible using suitable models and data from successive censuses. African statisticians and demographers will need to experiment with some of these mathematical models to derive results which may be relevant to development planning in their countries.

(b) Census and Civil/Vital Registration record

15. Although civil and vital registration records by themselves may contain no migration data, they provide the basis for the estimation of natural increase between two successive censuses. By using the balancing equation, it is possible to estimate net migration as the difference between the two census populations for the same areal unit less the excess of births over deaths. As already stated, this approach is valid only if the population is closed or virtually closed to international migration.

16. Unfortunately in Africa, very few countries have reliable statistics of births and deaths and those who have reasonable statistics tend not to be closed to international migration.

Thus the use of the vital statistics method in conjunction with place of birth statistics from two censuses has yet to be tried out extensively in the African context.

(c) Census and Survey:

17. Data from census or register and survey can usually supplement each other. Detailed migration data which cannot be collected in censuses can be obtained from sample surveys. By its very nature, census questionnaires can contain only limited questions like place of birth, place of usual residence, and place of usual residence x years ago. The more detailed questions such as place of previous residence, duration of residence, year of last move and reasons for migration are best collected in sample surveys.

18. In order to link the survey results with those of the census, three approaches on the timing of the survey may be used. The first is the so-called built-in sample which is convassed at the same time as the census. The sample units may be Enumeration Areas or households. The advantage of this approach is that by controlling on age group, sex and relationship to head of household, it is possible to use the ratio method of estimation to obtain estimates of characteristics collected in the survey with reduced variance. With all ratio estimators, there is some bias which can be regarded as neglible depending on the sampling fraction. The disadvantage of this approach is that for most African countries it is rather complex to administer, thus leading to significant non-sampling errors.

19. Difficulties that have been encountered in the administration of built-in samples in Africa include confusion between selected and non-selected units, namely between those units to receive the extended questionnaire and those also to be administered with only the short questionnaire, difficulties over identification and coverage of selected households and the cluster effect of taking whole primary sampling units such as Enumeration Areas as also the ultimate sampling units. However, in spite of these problems, ECA will continue to support this approach as one of the viable options open to countries that need district or sub-district estimates from their sampling to collect additional information as part of their overall population census programme.

20. The second approach is to have the survey three to six months after the census but to use the same reference date as the census. The estimation procedure is similar to the first. The main advantage is that because it is administered separately from the census and only after an interval of time, the complexity of its administration is reduced. The approach has however the disadvantage that the use of the census reference date leads to errors due to recall lapse.

21. The third approach is to have an independent survey after the census. Data collected in the census could be used as a basis for constructing a sample frame. The sampling units could be stratified on the basis of proportion of in-migrants in each Enumeration Area. In this context, it should be noted that the survey could be a general demographic survey with a migration component, a multi-subject survey with migration or migration related questions or a labour force survey with similar questions. The utilisation of data from the census or survey can sometimes be assisted by the matching of records from the two sources. This approach will be elaborated upon further in the section dealing with interlinking of survey and survey data.

22. The previous reference to the census providing a suitable frame for the follow-up migration survey needs further elaboration. Migrants (internal and international) are not evenly distributed among all primary sampling units (PSO). The census can thus be regarded as the first phase of a two-phase sampling scheme, where the first phase which consists of a complete enumeration enables stratification of p.s.u's on the basis of proportion of migrants. This is especially important for surveys which cover also international migration. Even for countries with relatively large proportions of internal migrants (say, more than 5 per cent of total population), these are concentrated in a few areas. This is even true of large cities, where migrants tend to be concentrated in a few areas. Thus an optimal sampling scheme will necessarily involve stratification based on information collected during the first phase survey (or census).

23. In case, the census and survey can be linked by any of the methods described above or by similar approaches, it is then possible to link the detailed data collected in the sample with the census data to obtain a more complete picture of the migration situation in a country. For example, if the census enumeration is on the basis of the de jure population, cross-classification of the population by region of enumeration (place of usual residence) and by region of birth will show persons who have moved since their births. The more detailed questions in the survey may then interpret these results by showing year of last move, region of previous residence, motivation for migrating etc.

(d) Census and population registers.

24. In Africa only a few countries have registration systems which even remotely resemble the population registers which exist in Scandinavian, some West European and other developed countries. Two examples of quasi-population registers are the records maintained by the kebele (Urban Dwellers Association) and the Peasant Associations in Ethiopia and the records of the Village Management Teams in Tanzania. In the former, all residents of the area are recorded and no migration even for a

short distance can be undertaken unless permission is granted. Thus with the use of the data generated from the system, it is possible to estimate changes in the population of an area since the previous census. There are however problems in making full use of the information so collected. The information is available at the level of kebele or Peasant Association. It is not aggregated at the Awraja (district) or wereda (region) level. Thus to try to work out migration from one region to another since the previous census on the basis of the records of the kebele and Peasant Association is a time-consuming operation. However this is a potential source of linking the stock data from censuses with the flow data from these continuous population registers.

25. Another problem with using this type of linkage in Africa is the reliability and accuracy of the records so maintained. By its very nature, a good population register should be continually updated, adding births, immigrants and in-migrants to the area and deleting deaths, emigrants and out-migrants. This is not only a time-consuming but also a costly operation. Very few African countries have the resources to maintain such registers and so at this time therefore they can only be regarded as potential sources of migration data. It is important to ensure however that when Governments set up such registers for their own internal administration, statisticians should get involved to assist in designing the system in such a way that their needs are taken into account i.e. that births, deaths and migrations are carefully recorded and that the completeness of the system for any area can be checked by the stock data at two successive censuses.

(e) Survey and Survey

26. Surveys repeated within a short-interval (not more than five years) also provide a basis for interlinking data from two identical sources. In the simplest example, a multi-round survey in which certain questions are repeated provides a good basis for interlinking data. This can be done firstly to verify the accuracy of respondents' responses during a previous round and secondly, to find changes in the migration histories of respondents which have taken place during the interval. This approach presupposes that records of subsequent rounds can be linked to those of the previous round and that the attrition rate (i.e. in this case, the proportion of respondents covered in the previous round who are no longer included in the subsequent round for any reason other than migration, e.g. death and non-responses) is high.

27. Multi-round demographic surveys were very popular in Francophone Africa in the late 1960's and early 1970's. However, the enthusiasm for this approach appears to have waned in the 1980's. Basically, the approach involves administering a base line survey in round one during which information on all usual members of

household is collected. Such information as shown in Annex I, which is taken from "Technical Problems of Multi-round Surveys" by C. Scott, (Laboratories for Population Statistics, 1973) could include the following items: for all household members: name, whether present or temporary absent, relationship to head of household, sex, age last birthday; for all women aged 14 years and over: total children born alive, still living; live births in last 12 months: number of live births, date of birth, whether baby is still living; if no, date of death, is woman pregnant. Round 2 of the multiround survey can be used to provide an update of the information collected during round 1. In addition new members since round 1 as well as deaths, births and migrations since round 1 can be recorded. Subsequent rounds follow the round 2 approach.

28. In the example given above, the follow-up survey is on births, deaths and migration. In case, the follow-up survey is on migration only, all references to births and dearhs in the second and subsequent rounds can be deleted. Instead, in addition to new members since round I and for those no longer resident who are not dead, information can be collected on residence now, place of previous residence, duration of stay at present residence, duration of stay at previous residence, place of usual residence 5 years ago and reason for leaving previous residence. Optionally, if this information cannot be derived from place of usual residence, information could be sought on distance between present and previous places of residence. Also, since socio-economic factors play an important role in interpreting the size and direction of migration information on marital status, educational attainment, type of economic activity, occupation, industry and employment status should be collected.

29. In the above example we have examined the case of respondents being linked when they are in different rounds of multi-round surveys. However, multi-round surveys in Africa have tended to be very complex, expensive and in certain countries have had an unacceptable attrition rate. Thus, African countries tend to have single round retrospective surveys rather than multi-round surveys. An examination of possiblities of interlinking data between two different single round retrospective surveys is therefore necessary as a means of assisting African countries to fully exploit migration data that may be collected in these surveys.

30. For most African countries such interlinking is possible at the macro-level, i.e. regional or district level. Based on birth place and usual residence data, estimates on net interregional or inter-districts migration can be obtained. If questions on place of residence one year ago are included in surveys which take place at intervals of one year, it is possible to compare the results obtained from the direct question with those obtained by estimating the net migraton during the preceeding year, through

indirect techniques.

31. Interlinking of responses in respect of the same household or individual is also possible even if only a part of the sample is retained for the subsequent round(s). A rotation sample is an example of such a design, in which, say, a fourth of the sample is replaced during each round so that at the end of other four rounds, the original sample has been completely replaced. The diagram below shows schematically what happens in each round:

1st round

A	B
D	C

2nd round

R1	B
D	C

3rd round

R1	R2
D	C

4th round

R1	R2
D	R3

5th round

R1	R2
R4	R3

where A, B, C, D represent the four subsamples and R1, R2, R3, R4 their replacements respectively.

32. In the case of this arrangement, data for B, C and D in the first and second rounds can be linked at the household and individual level. Similarly data from C and D can be linked at the first, second and third rounds, and data from D can be linked for four rounds.

33. Other sample designs are possible but will not be discussed in this paper which is merely meant to outline procedures for interlinking survey data at different rounds. Surveys held within the framework of the National Household Survey Capability Programme (NHSCP) and the Social Dimensions of Adjustment (SDA) can incorporate elements of linkages of migration data from the different rounds.

3. INTERNATIONAL MIGRATION

34. Section 2 of this paper has dealt primarily with internal migration. However not all these procedures are applicable to international migrants. A brief account of what can be done for international migrants is given below:

(a) Census and census

Comparisons can be made between stock data for the two censuses but information on international migration flows will require additional data on year of entry of the migrants enumerated in the second census as well as information of deaths among the international migrants enumerated in the first census. Since the latter is usually unavailable, the two censuses cannot generally be used together to obtain information which is more comprehensive than that obtained from either census.

(b) Census and population registers

Where population registers cover migrants and their particulars, it is possible to link the persons enumerated in the census with information in the registers to find out how many immigrants have died since the census, how many have left the country and how many have entered the country to stay. Thus at any given time, estimates of immigrants in the country can be made. Estimating the number of emigrants on the basis of census and population registers is more difficult and in the African situation almost impossible. Some African countries have during their censuses tried to estimate the number of emigrants by asking questions from heads of household members who have emigrated. This usually leads to gross underestimation of emigrants since households that have emigrated are left out. Thus a linkage between the census and register if it exists will at best reveal the weakness of this approach of estimating emigrants on the basis of the information supplied by heads of households.

(c) Census and survey

As in the case of internal migration,the census can provide information on the concentratiions of immigrants which could be used to stratify the primary sampling units for the purpose of collecting more detailed information on immigrants. The census could also be used as the basis of tracer survey where a sample of the persons enumerated in the census as immigrants could be followed and a more detailed migration questionnaire administered to them.

(d) Survey and survey

The same procedures outlined under section 2(a) apply here. The main problem is the estimation of the number of emigrants through linkages at the individual level. This can be overcome by the inclusion of additional questions which will assist in interpreting the results.

4. DUAL RECORD SYSTEM

35. A separate section is being devoted to this topic since its application covers both internal and international migration and since there are difficult conceptual and technical problems which are still the subject of controversy. The dual record system implies matching records of events (e.g. births, deaths, migration) on an individual basis and obtaining a better estimate of the number of events by applying the Chandrasekar - Deming formula. In the case of migration, the formula could be put as follows.

$$\hat{N} = \frac{\hat{N1}\ \hat{N2}}{M}$$

where

$\hat{N}$ = Chandrasekar - Deming estimate of the number of immigrants (or emigrants)

$\hat{N1}$ = Number of immigrants (or emigrants) from first source

$\hat{N2}$ = Number of immigrants (or emigrants) from second source

M = Number of immigrants who could be matched (i.e. the number of immigrants appearing in both sources).

36. The sources could be census, population registers or survey. One requirement for the application of the method is the statistical independence of the two sources, which is often difficult to meet since, for example, a survey is based on a frame provided by the census. There are also problems of matching such as lack of uniqueness of names and ambiguities in the identification of housing units to which the persons to be matched belong(ed).

37. In spite of these problems, this is an area for further experimentation in which the theoretical basis of the application of the Chandrasekar-Deming formula has to be re-examined and the estimates obtained subjected to intensive analysis.

CONCLUSION

38. The purpose of the paper is to suggest ways of improving the quality of migration data in Africa and to outline for further research. Model building has now become an accepted part of attempts to interprete levels and trends in the components of population growth. African statisticians and analysts need to familiarize themselves with the latest techniques in this area and to carry out experimenrs which will improve the quality of migration data in Africa. Most of the approaches suggested in this paper have yet to be extensively explored in the region but as an outcome of this meeting it is hoped that both data producers and users will document their experiences in the use of these methods and such documentation will be available to all interested bodies.

Questionnaire I. Draft questionnaire for follow-up survey on births deaths and migration*

Region.....
Distr.....
Town or village .....
Street name & No. if any.....
E.A. No.....
Household.....

Country and date
PART A
See reverse for part B

FOR ALL USUAL MEMBERS OF HOUSEHOLD						FOR ALL WOMEN AGED 14 Yrs. OR MORE							NO LONGER RESIDENT AT ROUND 2		NO LONGER RESIDEN AT ROUND 3	
						TOTAL CHILDREN		LIVE BIRTHS IN LAST 12 MONS.								
Serial No.	Name	Present or temporarily absent P A	Relationship to head of household	Sex M or F	Age last birth-day	Born alive M : F	Still living M : F	No. of live births M : F	Date of birth	Baby still living Yes No	If No. date of death	Is this woman pregnant now Yes No	If dead date of death	If living elsewhere residence now	If dead date of death	If living elsewhere residence now
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	18	19
ROUND 1									Date of visit.....				Date of visit		Date of visit	
1						:	:	:								
2						:	:	:								
3						:	:	:								
4						:	:	:								
5						:	:	:								
6						:	:	:								
7						:	:	:								
8						:	:	:								
9						:	:	:								
10						:	:	:								
11						:	:	:								
12						:	:	:								
13						:	:	:								
14						:	:	:								
15						:	:	:								
ROUND 2 NEW MEMBERS SINCE ROUND 1													For babies under 1 yr. serial no. of mother	For all arrivals from elsewhere. previous residence	: :	: :
1	2	3	4	5	6	7	8	9	10	11	12	13	16	17	18	19
21						:	:	:								
22						:	:	:								
23						:	:	:								
24						:	:	:								
25						:	:	:								

* Reproduced from "Technical Problems of Multi-round Demographic Surveys" Laboratories for Population Statistics Report series No. 11 (1973).

Questionnaire I (continued)

Questionnaire I (continued)

ROUND 3 NEW MEMBERS SINCE ROUND 2

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Remarks
													For babies under 1 yr. of mother: serial no.	For all arrivals from elsewhere. previous residence	
31															
32															
33															
34															
35															

Suggested definitions

A household consists of one person living alone, or of a group of persons living together and having common provisions for eating

A usual resident is one who has lived with the household for 6 months or more, or intends to remain 6 months or more

Temporary means for less than 6 months.

A live birth means the birth of a baby who cried or breathed, if only for a moment.

PART B

LIVE BIRTHS AND DEATHS IN HOUSEHOLD IN LAST 12 MONTHS

N.B. Cols. 9 to 12 should include all deaths, including any reported in part A and any removed in cols. 5 and 6 of part B'

LIVE BIRTHS								DEATHS			
						For women not listed in Part A					
Serial No.	Sex M F	Date of birth	Baby still living	If dead date of death	Serial no. of women in part A	If living elsewhere, place of residence	If dead, age at death	Serial No.	Sex M or F	Date of death	Age at death
1	2	3	4	5	6	7	8	9	10	11	12
ROUND 1											
ROUND 2											
ROUND 3											